Reprinted from

Canadian Journal of Fisheries and Aquatic Sciences

Réimpression du

Journal canadien des sciences halieutiques et aquatiques

Evidence of Contaminant Loading to Lake Ontario from the Niagara River

A. H. El-Shaarawi, S. R. Esterby, N. D. Warry, and K. W. Kuntz

Volume 42 • Number 7 • 1985

Pages 1278–1289

Canada

Fisheries and Oceans Pêches et Océans

Printed in Canada by University of Toronto Press

Evidence of Contaminant Loading to Lake Ontario from the Niagara River

A. H. El-Shaarawi

Aquatic Physics and Systems Division, National Water Research Institute, Burlington, Ont. L7R 4A6

S. R. Esterby

Aquatic Ecology Division, National Water Research Institute, Burlington, Ont. L7R 4A6

and N. D. Warry and K. W. Kuntz

Water Quality Branch, Inland Waters Directorate – Ontario Region, Burlington, Ont. L7R 4A6

El-Shaarawi, A. H., S. R. Esterby, N. D. Warry, and K. W. Kuntz. 1985. Evidence of contaminant loading to Lake Ontario from the Niagara River. Can. J. Fish. Aquat. Sci. 42: 1278–1289.

The hypothesis that the concentrations of pollutants are the same throughout the length of the Niagara River, against the alternative that they are higher in the lower river, was examined using three sets of data: (1) Ontario Ministry of Environment's (MOE) results for organic substances in raw drinking water (1978–84); (2) MOE transect data for metals in water (1981); (3) Environment Canada data on organic substances in large-volume water samples and suspended sediments (1981). The designs of the studies dictated different statistical procedures for each set, with nonparametric procedures based on ranking being used for sets 1 and 3, and normal theory methods for set 2. Further, both univariate and multivariate procedures were used for set 3. Despite the considerable number of nondetections in these data, it was still possible to make quantitative comparative statements, with the degree of uncertainty providing the quantitative component. In each study, a number of substances were shown to be higher in the lower river than in the upper river, and by examining all substances of set 3 simultaneously, locations in the lower river were shown to be similar to each other as were locations in the upper river.

L'alternative hypothétique suivante : les concentrations de matières polluantes sont semblables dans tout le cours de la rivière Niagara, ou les concentrations sont plus élevées dans le cours inférieur, a été évaluée à l'aide de trois séries de données : (1) les résultats de quantification des substances organiques dans l'eau potable non traitée recueillis par le ministère ontarien de l'Environnement (MOE) de 1978 à 1984; (2) les données par coupe transversale sur les métaux dissous dans l'eau, recueillies par le MOE en 1981; (3) les données d'Environnement Canada sur les substances organiques contenues dans de grands volumes d'eau échantillonnés et dans les sédiments en suspension (1981). La conception des études a déterminé le genre d'analyse statistique appliquée à chaque série de données : des méthodes non paramétriques basées sur le classement ont été utilisées pour les séries 1 et 3 et la méthode basée sur la théorie normale, pour la série 2. De plus, la série 3 a fait l'objet d'analyses unifactorielle et multifactorielle. Malgré le nombre élevé de non-détections dans ces données, on a pu faire des comparaisons quantitatives où le degré d'incertitude représentait la composante quantitative. Dans chaque étude, la concentration d'un certain nombre de substances s'est révélée plus élevée dans le cours inférieur que dans le cours supérieur de la rivière. L'examen simultané de toutes les substances de la série 3 a démontré que les emplacements dans le cours inférieur étaient semblables, de même que ceux dans le cours supérieur.

Received September 27, 1984
Accepted April 4, 1985
(J7968)

Reçu le 27 septembre 1984
Accepté le 4 avril 1985

Numerous studies have demonsurated that the Niagara River is a major source of the pollutants that enter Lake Ontario. The problem, summarized in Allan et al. (1983), is that the Niagara River, which is the only outflow for Lake Erie and the major tributary to Lake Ontario (83% of the total water input), receives chemical substances from waste dumps and municipal and industrial outfalls, primarily from the concentration of chemical industries on the New York State side of the river.

Although it is clear that chemicals are entering Lake Ontario from the Niagara River, it is difficult to make quantitative comparative statements about the amounts of many organic substances entering the lake because environmental concentrations are too close to the analytical limits of detection. We address the question, are the concentrations of substances in water and sediments higher in the lower river than in the upper river? Three very different sets of data are used, and thus different statistical methods are applied. The results show that

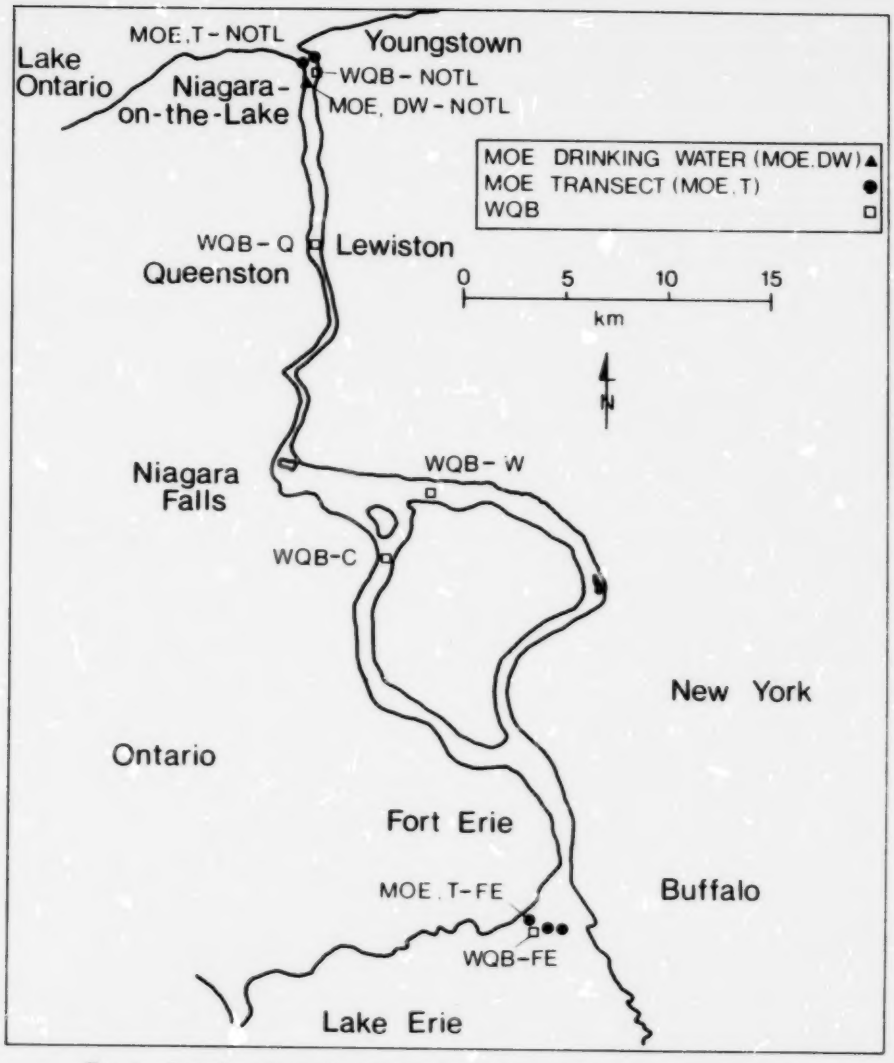

FIG. 1. Location of sampling stations on the Niagara River for all three studies.

comparative statements can be made quantitative, where the quantification is given by the degree of uncertainty (usually the significance level) attached to the statement, and that this can be done even when a substantial number of nondetections occur in the data.

Data and Methods

Data Sets

The following three data sets, collected under different sampling strategies, were available for analysis:

(1) From the Ontario Ministry of the Environment (MOE) monitoring program for volatile organics in raw drinking water, the results of the analyses for 24 of the substances classified as priority pollutants by the US EPA, 22 other organics, and 23 pesticides in water samples collected at Fort Erie (FE) and Niagara-on-the-Lake (NOTL) between August 1978 and January 1984.

(2) From the MOE transect study, the concentrations of nine metals in samples from the three stations on the FE transect which are representative of Lake Erie and all five stations on the NOTL transect, where sampling was done on three successive days in the months of June, August, and September of 1981.

(3) The results of the analysis for 28 chlorinated organic substances in large-volume water samples and in suspended sediments obtained by the Water Quality Branch (WQB), Environment Canada, at five stations (Fig. 1): FE, Chippawa (C), Wheatfield (W), Queenston (Q), and NOTL, with one or two samples per station once during May, July, and October–November of 1981.

In contrast with the last two sets of data, drinking water was collected over a number of years, at various times during the year, and with different frequencies for the two stations. Not all substances were reported for each sample, and the result reported was often a nondetection. The metal data was mostly quantitative, with some metals not reported for some samples and some results removed because they had been flagged as being of questionable quality. The stations used within a transect are not expected to show spatial variability. In the WQB study, results were reported for all substances in all samples, but these included numerous nondetections. Since data for only two

locations were used from the MOE data sets, the objective is to determine if there are differences in the levels of a substance between the locations. The data from five locations in the WQB study permit not only the analogous comparison of levels, but also the analysis for trend due to position down the river.

Analytical Methods

The analytical methods used for these three data sets have been independently reviewed and accepted by the Canada – United States Niagara River Toxics Committee (Data Quality Subcommittee 1984). In the MOE studies, pesticides and PCBs were analyzed using gas chromatography with an electron capture detector to detection limits of 0.001–0.005 μg/L for pesticides and 0.02 μg/L for PCBs. GC/MS purgeables were analyzed to a detection limit of 0.1 μg/L and the GC/MS extractable analysis was qualitative only. Metals were determined using either flame atomic absorption or inductively coupled plasma spectrometry. Total tetrachloro dibenzo-*p*-dioxins were analyzed by GC/MS with a detection limit of 0.25 ng/L prior to May 1983, and after May 1983 to a detection limit for each individual isomer of tetrachloro dibenzo-*p*-dioxins of 0.01 ng/L (MOE 1984). The method used for the WQB large-volume water samples (McCrea 1982) and suspended sediments (Kuntz and Warry 1983) consists of collection of a sample over an approximate 2-d period during which the suspended sediment sample is separated by means of a Westfalia continuous-flow separator and a 200-L sample of the effluent is collected for the water sample. The water samples are extracted with dichloromethane, the sediment samples pressure-filtered, dried, and ground, and then both analyzed at the WQB Organic Analysis Laboratory according to methods described in the *Analytical Methods Manual* (Environment Canada 1979).

Statistical Methods

For the drinking water data, the comparison of the level of a substance at NOTL and the level at FE has been based on samples taken at the two stations at about the same time (usually the same day). The sign test (e.g. Lehmann and D'Abrera 1975) was applied if enough matched sampling dates were available for a substance. The sign for the pair was determined by ranking one station relative to the level of the other. Let (x_{i1}, x_{i2}) be the results for the analysis of the substance under study at locations 1 and 2 on day i for $i = 1, 2, \ldots, n$. Then

$d_i = +$ if either

(i) x_{i1} and x_{i2} are measurable and $x_{i1} > x_{i2}$,

(ii) x_{i1} is measurable and x_{i2} is recorded as trace (Tr) or not detected (ND), or

(iii) x_{i1} = Tr and x_{i2} = ND;

$d_i = -$ if either

(i) x_{i1} and x_{i2} are measurable and $x_{i1} < x_{i2}$,

(ii) x_{i1} = Tr or x_{i1} = ND and x_{i2} is measurable, or

(iii) x_{i1} = ND and x_{i2} = Tr;

and $d_i = 0$ if either

(i) x_{i1} and x_{i2} are measurable and $x_{i1} = x_{i2}$,

(ii) $x_{i1} = x_{i2}$ = Tr, or

(iii) $x_{i1} = x_{i2}$ = ND.

Let N_+, N_-, and N_0 be the number of d_i that are $+$, $-$, and 0, respectively. The objective is to test the hypothesis of no difference between the concentrations of a substance at the two stations against the alternative that the concentrations are higher at NOTL. In the presence of ties, $N_0 \neq 0$, two statistics are available (Lehmann and D'Abrera 1975). The most powerful test is based on the statistic N_+ (Putter 1955), which is obtained by conditioning on the observed number of ties, $N_0 = n_0$, and is equivalent to omitting the ties. A statistic that includes the ties is $N_+ + 1/2N_0$, but the distribution of this contains the unknown parameter, p_0, which is the probability of obtaining a tie. N_+ has a binomial distribution with parameters $n - n_0$ and $p = 0.5$, and the exact significance level associated with an observed value of N_+ can be calculated. The customary procedure for obtaining the significance level for $N_+ + 1/2N_0$ is to use the binomial distribution with parameters n and $p = 0.5$ (Putter 1955). The asymptotic test is the same for both statistics and is obtained from

$$Z = (2N_+ - n + N_0)/(n - N_0)^{1/2}$$

which has approximately a $N)0, 1)$ distribution.

The analysis of the transect data for metals at NOTL and FE was based on a partially nested model (Scheffé 1959). Days are nested within months but months and stations are crossed factors. The model is given by

$$x_{ijkl} = \mu + m_i + d_{j(i)} + s_k + (sm)_{ik} + (sd)_{jk(i)} + e_{ijkl}$$

where m_i is the effect of month i, $d_{j(i)}$ is the effect of day j within month i, s_k is the station effect for station k, $(sm)_{ik}$ is the interaction between month i and station k, and $(sd)_{jk(i)}$ is the interaction between station k and day j within month i. The subscript l refers to replicates and the e_{ijkl} are assumed independent $N(0, \sigma^2)$.

The design must either be balanced (equal number of replicates) or have proportional frequencies (Kendall and Stuart 1968) if the sums of squares are to be calculated in the usual way (Bennett and Franklin 1954) as sums of the squared differences between means. Because of the unequal number of replicates for the two stations, the sums of squares were obtained using "dummy" variables and regression methods (Draper and Smith 1981). The residual sum of squares obtained when all main effects and interactions are included in the model provides the error sum of squares for testing the significance of effects. The sums of squares of main effects and interactions are obtained individually as the difference between the residual sum of squares, with the parameters corresponding to the particular main effect or interaction omitted, and the residual sum of squares for the full model. Once the sums of squares are calculated the analysis of variance can be conducted in the usual way.

In the presence of significant interactions, the effects of a corresponding factor must be examined at each level of the other factor. The t-test was used to compare stations for the individual days. Let $\bar{x}_{ijk}$ and s^2_{ijk} be the mean and the variance from the n_{ijk} samples taken at station k on day j of month i, and s^2 be the within-day variance obtained by pooling a number of s^2_{ijk}. Then the t-test for day j, with $k = 1$ for NOTL and $k = 2$ for FE, is given by

$$t = (\bar{x}_{ij1} - \bar{x}_{ij2})/s \sqrt{\frac{1}{n_{ij1}} + \frac{1}{n_{ij2}}}$$

where t follows a t-distribution with degrees of freedom equal to that of s. For the days exhibiting higher variation, the test due to Welch (Brownlee 1965) for the equality of means in samples with different variances was applied. The test statistic is

$$t' = (\bar{x}_{ij1} - x_{ij2})/\sqrt{A + B}$$

where $A = s^2_{ijl}/n_{ijl}$, $B = s^2_{ij2}/n_{ij2}$, and t' has approximately a t-distribution with degrees of freedom

$$f' = \frac{(A + B)^2}{A^2/(n_{ij1} - 1) + B^2/(n_{ij2} - 1)}.$$

The statistical analysis of the WQB data attempts to answer the following questions: (1) Are there overall differences between the levels of the chlorinated organics at the five sampling stations? (2) If the answer is yes in (1), then, is there a consistent pattern for the differences? (3) What is the direction of this pattern? (4) What can be said about the behaviour of individual substances? The analysis was performed using ranks that were assigned as follows: let $x_1, x_2, \ldots, x_n$ be the concentrations of a specific substance at a given data in samples from locations 1, 2, ..., n, where x is the mean for those dates when two samples were taken. The smallest x_1 is replaced by 1, the second smallest by 2, ..., and the largest by n. In the case of ties, the mid-rank is used. Rank one is given when x is below the detection limit, and the mid-rank is used when there is more than one x below the detection limit. Thus, the data tables were replaced by tables of the ranks prior to the analysis.

Let R_{ij} be the rank of the ith substance ($i = 1, 2, \ldots, l$) at the jth station ($j = 1, 2, \ldots, n$) where l is the number of substances. The squared Euclidean distance, ED(j, j'), between locations j and j' is

$$\mathrm{ED}(j, j') = \sum_{i=1}^{l} (R_{ij} - R_{ij'})^2.$$

It can be shown that under the hypothesis of no differences between the stations and independence between the substances, the mean of ED(j, j') is

$$M = \frac{ln(n + 1)}{6} - \frac{\sum_{i=1}^{l} \sum_{s=1}^{g} d_{is}(d_{is}^2 - 1)}{6(n - 1)}$$

where for the ith substance d_{i1} gives the number of locations with the common rank $(d_{i1} + 1)/2$, d_{i2} gives the number of locations with the common rank $d_{i1} + (d_{i2} + 1)/2$, and so on. When there are no tied observations for the ith substance, all $d_{is} = 1$. The expression for the variance is

$$\sigma^2 = \frac{2}{n - 1} \sum_{i=1}^{l} \left\{ \sum_{s} U_{is}^4 - 2(n + 1) \sum_{s} U_{is}^3 + 3\left[\frac{n(n + 1)(2n + 1)}{6} - \frac{\sum_{s} d_{is}(d_{is}^2 - 1)}{12}\right]^2 \right\} - M^2$$

where $U_{is} = d_{i1} + d_{i2} + \ldots + d_{is} + d_{is+1}/2$. In the case where $d_{is} = 1$ for all i and s, the above expressions reduce to

$$M = \frac{ln(n + 1)}{6}$$

and

$$\sigma^2 = ln(n + 1)(7n^2 - 5n - 18)/180.$$

The standardized Euclidean distance is

$$\mathrm{SED}(j, j') = \{\mathrm{ED}(j, j') - M\}/\sigma.$$

Under the assumptions that the locations are equally similar to each other, the values of SED(j, j') will be close to zero. Large negative (positive) values of SED(j, j') indicate that the locations are more similar (less similar) than that expected under equal similarity between the locations. The values of SED(j, j') are used to form clusters using the complete-linkage clustering method (Anderberg 1973).

The differences between locations for each individual substance are examined as follows. Let x_{ij} be the concentration of the substance of interest at the ith time ($i = 1, 2, \ldots, a$) and the jth location ($j = 1, 2, \ldots, n$) and let R_{ij} be the rank of x_{ij} in the joint ranking of $x_{i1}, x_{i2}, \ldots, x_{in}$. The Page test (Hollander and Wolfe 1973, p. 147) is designed to test for an increasing trend. The test statistic is

$$S_1 = \sum_{j=1}^{a} jR_j$$

where

$$R_j = \sum_{i=1}^{n} R_{ij}.$$

Critical values of S_1 are given for small values of a and n in table A(16) in Hollander and Wolfe (1973). For large n the distribution of S_1 is normal with mean $na(a + 1)^2/4$ and variance $n(a^3 - a)^2/144(a - 1)$, and hence the normal distribution could be used to calculate the significance level associated with an observed value of S_1.

Results

MOE Drinking Water Data

The hypothesis of equal concentration at NOTL and FE versus the alternative that the concentration is higher at NOTL than at FE was tested for 15 of the priority pollutants, eight of the organic substances, and two of the pesticides using the one-sided sign test (Table 1). The test was performed only if the number of pairs, n_d, exceeded three, since even if the NOTL level is higher than the FE level for each pair when the number of pairs of samples equals three, the significance level is 0.13. Many of the 69 substances were either never detected or detected no more than three times (Tables 2 and 3). Of the 25 substances tested, the following four were significant at the 1% level: chloroform, trichloroethylene, tetrahydrofuran, and *p*-chlorotrifluorotoluene (Table 1). In addition, two more were significant at the 5% level (tetrachloroethylene and A-BHC), and three more at the 10% level (benzene, 2,4,6-trichlorophenol, and chlorotoluene) for a total of four substances significant at the 1% level, six at the 5% level, and nine at the 10% level. These conclusions hold for the tests based on both N_+ and $N_+ + 1/2N_0$, and over all the substances, the two significance levels were very similar. Note that the significance level for dimethyldisulphide is high, since all of the six samples showed higher levels at FE than at NOTL and the one-sided test conducted here gives the probability of obtaining a number of positive signs at least as large as that observed, namely $r = 0$. That is, the only evidence included, against the hypothesis of equal concentration, is evidence in the direction of NOTL > FE.

A cumulative sum plot can be used to examine the constancy of the difference between the stations over time. For a particular substance, let R_j = the number of times the level of the substance at NOTL exceeds that at FE for the first j dates. The plot of the cumulative sum of the difference between R and its expected value $E(R)$ is shown for chloroform in Fig. 2. For the

TABLE 1. Test of the hypothesis of equal concentration at NOTL and FE for substances in raw drinking water based on matched sampling dates, MOE data.

Class	Compound	n_+	$n - n_0$	Pc	n_0	$n_+ + 1/2n_0$	P
Priority pollutants	Methylene chloride	2	4	0.69	0	0	Pc
	Chloroform	30	43	0.007	10	35	Pc
	1,1,1-Trichloroethane	4	6	0.34	3	5.5	0.26
	Dichlorobromomethane	10	16	0.23	0	10	Pc
	Trichloroethylene	15	18	0.004	2	16	0.006
	Chlorodibromomethane	11	17	0.17	0	11	Pc
	Benzene	16	24	0.08	12	22	Pc
	Tetrachloroethylene	8	10	0.05	3	9.5	0.05
	Toluene	10	24	0.85	13	16.5	Pc
	m-Xylene	5	9	0.50	4	7	0.50
	o,p-Xylene	8	13	0.29	6	11	0.32
	Hexane	9	13	0.13	4	11	0.17
	Hexanal	5	9	0.50	4	7	0.50
	Tetrahydrofuran	23	23	0	0	23	Pc
	p-Chlorotrifluorotoluene	10	10	0.001	0	10	Pc
Organics	Dimethyldisulphide	0	6	1	1	0.5	0.99
	Diethyl ether	5	7	0.23	0	5	Pc
	2,4,6-Trichlorophenol	4	4	0.06	0	4	Pc
	Carbon disulphide	3	4	0.31	1	3.5	0.19
	Methyl ethyl ketone	4	7	0.50	2	5	0.50
	Chlorotoluene	4	4	0.06	0	4	Pc
	Acetone	5	9	0.50	11	10.5	0.42
	Hydrocarbons	8	16	0.60	10	13	0.55
Pesticides	A-BHC	18	26	0.04	6	21	Pc
	G-BHC	6	9	0.25	2	7	0.27

NOTE: n_+ = number of times the level was higher at NOTL than at FE. n_0 = number of times the levels were the same at NOTL and FE. n = total number of pairs. Pc = one-sided observed significance level, based on n_+. It is the exact significance level if $n - n_0 \leq 20$ or the normal approximation if $n - n_0 > 20$. P = one-sided observed significance level based on $n_+ + 1/2n_0$. It is calculated from the binomial distribution with parameters n and $1/2$ if $n \leq 20$ or the normal approximation if $n > 20$. To calculate the significance level from noninteger values and the binomial distribution, the smallest integer $> n_+ + 1/2n_0$ was used. Pc appearing in this column indicates that the significance level equals that based on n_+.

TABLE 2. Compounds not detected at either NOTL or FE on matched sampling dates, MOE drinking water data.

Priority pollutants	Organics	Pesticides
Dichloroiodomethane	Total trichlorophenols	Aldrin
1,1,4-Trichlorobutadiene	2,3,4,5-Trichlorophenol	DMDT–methoxychlor
Bromoform	3,4,6-Trichlorophenol	Endosulphan II
	Dichloropropylene	Endosulphan sulphate
		Heptachlor
		Mirex
		Oxychlordane
		o,p-DDT
		p,p-DDD
		p,p-DDE
		Total tetrachlorodioxins

*i*th sampling date this is

$$\sum_{j=1}^{i} (R_j - E(R_j)) = \sum_{j=1}^{i} (R_j - j/2)$$

and this has been plotted against the date, to the nearest half-month, which corresponds to i. If there is no difference between the chloroform concentration at NOTL and FE, the points in Fig. 2 should fluctuate around zero. On the contrary, the points are always above zero and continue to increase throughout most of the sampling period. This indicates that the concentration at NOTL remains above that at FE, except possibly from about July 1981 to the end of 1982.

MOE Transect Data

The analysis of variance was performed for four metals (Table 4). The interactions between station and day within month and between station and month were significantly differ-

TABLE 3. Compounds detected not more than three times at either NOTL or FE on matched sampling dates, MOE drinking water data.

Priority pollutants			Organics			Pesticides		
	Occurrences			Occurrences			Occurrences	
Compound	NOTL	FE	Compound	NOTL	FE	Compound	NOTL	FE
1,1-Dichloroethane	1	0	2,3,5-Trichlorophenol	2	0	A-chlordane	1	0
1,2,-Dichloroethane	1	0	2,3,6-Trichlorophenol	1	0	G-chlordane	1	0
1,2-Dichloropropane	1	0	Dichloroethylene	1	0	Endosulphan I	1	0
Chlorobenzene	1	0	Dichlorobenzene	3	0	Endrin	1	0
Ethyl benzene	2	0	Styrene	1	0	*p,p*-DDT	0	1
Carbon tetrachloride	2	1	Heptanone	2	0	Hexachlorobenzene	2	0
			Benzaldehyde	0	1	Total PCB	2	0
			Methyl furan	0	3	Dieldrin	0	2
			Trichlorotrifluoroethane	0	2	Heptachlor expoxide	0	2
			Butanol	2	0	B-BHC	3	0

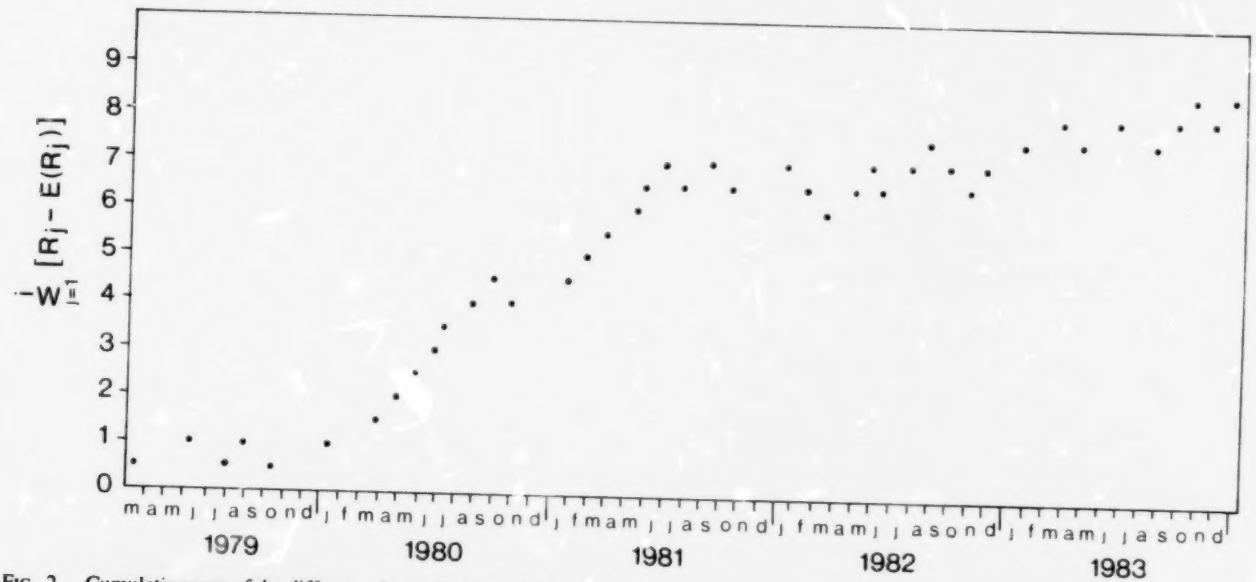
$-W\sum_{j=1}^{i}[R_j - E(R_j)]$
9
8
7
6
5
4
3
2
1
0
m a m j j a s o n d j f m a m j j a s o n d j f m a m j j a s o n d j f m a m j j a s o n d j f m a m j j a s o n d
1979
1980
1981
1982
1983

FIG. 2. Cumulative sum of the difference between the number of +'s up to sample i and its expected value calculated for MOE chloroform data and plotted against the date of sample i.

TABLE 4. Analysis of variance based on partially nested model for four of the metals in the MOE transect study (df = degrees of freedom, MS = sum of squares, F = ratio of mean squares, P = observed significance level).

DF (replicates, total) source	df	Iron (53,70)			Aluminum (53,70)			Zinc (52,69)			Copper (51,68)		
		MS	F	P	MS	F	P	MS	F	P	MS	F	P
Between months	2	0.00495	2.86	0.06	0.00021	1.30	0.272	6.47×10^{-6}	0.20	0.975	1.23×10^{-5}	0.09	0.996
Between days within months	6	0.00290	1.68	0.14	0.00452	9.35	<0.001	7.94×10^{-5}	2.41	0.098	0.00078	5.87	0.005
Between stations	1	0.00306	1.77	0.19	0.00017	0.18	0.677	1.33×10^{-7}	0.00	0.905	7.01×10^{-6}	0.05	0.805
Station × month interaction	2	0.00809	4.68	0.01	0.00072	4.48	0.001	1.74×10^{-6}	0.05	0.999	2.03×10^{-5}	0.15	0.987
Station × day within month interaction	6	0.00470	2.72	0.02	0.00594	12.29	<0.001	8.66×10^{-5}	2.63	0.080	0.00066	4.90	0.011
Between replicates		0.00173			0.00097			3.29×10^{-5}			0.00013		
Total (corrected)		0.00610			0.00324			4.02×10^{-5}			0.00015		

TABLE 5. Summary of analyses of MOE transect metal data where units of concentration are mg/L.

		Iron					Aluminum					Zinc				
		Mean concentration		Significance level			Mean concentration		Significance level			Mean concentration		Significance level		
Month	Day	NOTL	FE	s^2 9 days	s^2 7 days	$s_1^2 \neq s_2^2$	NOTL	FE	s^2 9 days	s^2 7 days	$s_1^2 \neq s_2^2$	NOTL	FE	s^2 9 days	s^2 8 days	$s_1^2 \neq s_2^2$
6	11	0.310	0.129	<0.0001	<0.0001		0.236	0.133	<0.0001	<0.0001		0.014	0.013	0.41	0.36	
	12	0.176	0.096	0.005	0.001		0.160	0.100	0.005	0.0006		0.012	0.007	0.13	0.03	
	13	0.160	0.054	0.0005	—	0.02	0.152	0.075	0.008	—	0.04	0.002	0.005	0.73	0.84	
8	5	0.067	0.088	0.75	0.81		0.084	0.095	0.69	0.74		0.005	0.006	0.62	0.68	
	6	0.112	0.052	0.03	0.008		0.107	0.069	0.05	0.015		0.010	0.002	0.03	—	0.18
	7	0.110	0.125	0.80	—	0.65	0.111	0.057	0.01	0.002		0.005	0.008	0.72	0.83	
9	23	0.222	0.076	<0.0001	<0.0001		0.206	0.124	0.001	—	0.17	0.004	0.003	0.40	0.35	
	24	0.170	0.075	0.0015	0.0002		0.180	0.102	0.0007	<0.0001		0.005	0.004	0.37	0.30	
	25	0.096	0.056	0.09	0.05		0.096	0.086	0.34	0.29		0.003	0.003	0.45	0.44	

		Copper					Mean concentration							
		Mean concentration		Significance level			Nickel		Chromium		Cadmium		Lead	
Month	Day	NOTL	FE	s^2 9 days	s^2 7 days	$s_1^2 \neq s_2^2$	NOTL	FE	NOTL	FE	NOTL	FE	NOTL	FE
6	11	0.011	0.018	0.80	1.00		0.003	0.004	0.002	0.002	0.002	0.002	0.010	0.004
	12	0.020	0.011	0.15	—	0.25	0.002	0.002			0.002	0.005		
	13	0.001	0.011	0.84	1.00		0.005	0.002	0.004	0.026	0.003	0.002		
8	5	0.003	0.007	0.67	0.89		0.003	0.006	0.003	0.002	0.003	0.002	0.005	0.005
	6	0.008	0.003	0.30	0.07		0.003	0.003	0.003	0.002	0.002	0.003	0.003	0.005
	7	0.007	0.007	0.54	0.57		0.002	0.003	0.002	0.002				
9	23	0.003	0.006	0.61	0.78		0.002	0.003	0.004	0.009				
	24	0.005	0.032	1.00	—	0.82	0.005	0.002	0.004	0.002	0.003	0.007		
	25	0.006	0.004	0.40	0.73		0.002	0.004	0.003	0.002	0.003	0.003	0.005	0.006

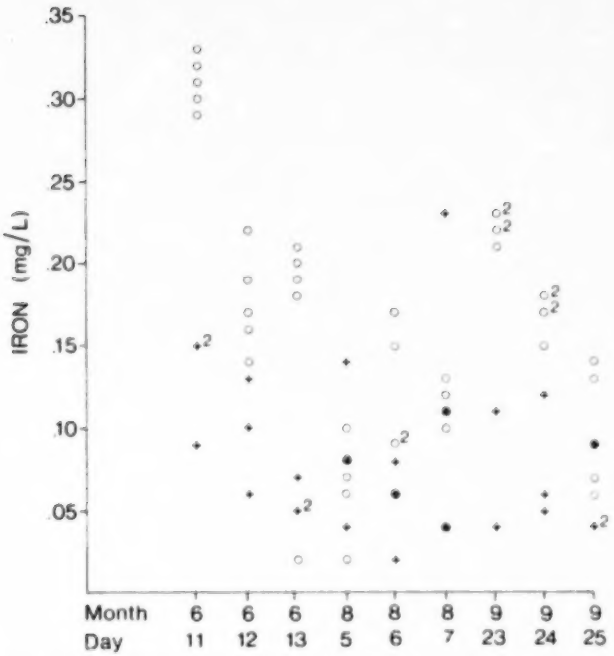

FIG. 3. Iron concentration by day for NOTL (○) and FE (+) stations, MOE transect data. The numbers beside the symbols indicate the number of stations with the same concentration within the resolution of the plot.

ent from zero at the 0.02 level or lower for iron and aluminum, and between station and day within month at the 0.01 level for copper. No effects were significantly different from zero for zinc. The analysis could end here for zinc. Due to the significant interactions, the station effect for iron, aluminum, and copper had to be evaluated for each day (Table 5).

Logarithms are often used in the analysis of environmental data. Since replicates were available, the relationship between the variance and mean was examined by plots of s^2_{ijk} versus $\bar{x}_{ijk}$. The increase in s_{ijk} with $\bar{x}_{ijk}$, which motivates the transformation to logarithms of the concentration, was definitely not present for this data set. The major feature was one or two larger within-day variances. The residual sum of squares in the present analysis of variance is the between replicate sum of squares which is given by

$$\sum_i \sum_j \sum_k \sum_l (x_{ijkl} - \bar{x}_{ijk})^2 = \sum_i \sum_j \sum_k (n_{ijk} - 1)s^2_{ijk}.$$

The residual mean square provides the estimate of the variance to be used to test equality of station means separately for each day. Where larger values of s^2_{ijk} were indicated from the plots, the tests of station effect were also evaluated with those excluded. Thus, in Table 5, there are three columns of significance levels for each metal. The first two correspond to the result of a t-test using the estimate of the variance given by the between-replicates mean square (Table 4) and by the between-replicates mean square where days with larger variance at either NOTL or FE have been eliminated. These t-tests are based on degrees of freedom in the range 39–53. For the days exhibiting larger variation, the Welch test was applied and these results appear as the third column in each set of columns of significance levels. Significance levels have been quoted for zinc, since weak evidence of interactions was present. Because the concentrations show very little variability and were reported to only one significant digit, it was decided that the analytical method used for the other five metals was not sensitive enough to detect differences in concentration. Thus, the above-described tests were not applied. Mercury was completely omitted because it was analysed on 3 d only.

Of the metals, only iron and aluminum had concentrations reported to two or three significant digits. Further, with the exception of zinc on August 6, only iron and aluminum showed significant differences between NOTL and FE. For iron, the concentrations at the two stations were more different in June and September than in August and for aluminum in June (e.g. Fig. 3).

WQB Data: Differences between Locations

Figures 4A and 4B showed that NOTL and Q are similar to each other (SED(Q, NOTL) < 0) and different from the rest of the locations. The locations FE, C, and W always form a group with SED(j, j') negative, except for the chlorinated organics in the aqueous phase in July where location W was linked with locations FE and C at a positive value for SED(j, j').

To determine the pattern of the differences between the locations, the substances were divided into three groups (PCBs, pesticides, and chlorinated aromatics). Under the assumption that there are no differences between locations, it is expected that the mean rank for each group should be close to three. The plots of mean ranks (Fig. 5A and 5B) show that (1) the pesticides are not good for differentiating between locations and (2) NOTL and Q are similar to each other and different from FE, C, and W. In fact, the mean ranks for NOTL and Q are consistently greater than three for the chlorinated aromatics and PCBs in both the suspended and the aqueous phase. This indicates an overall increase in the concentrations of PCBs and the chlorinated aromatics in the lower Niagara River as compared with the upper part of the river.

WQB Data: Trend Analysis for Individual Substances

Table 6 gives the value of S_1 for each substance, for the suspended sediments, aqueous phase, and for combined data. The results show the existence of a significant increase in the concentration of 16 substances in either the suspended phase or in the aqueous phase, while for six substances most of the data were below the detection level and hence the analysis was not performed. This shows that eight substances had no increasing trend from FE to NOTL.

Discussion

Each study provides evidence that some substances are present in either water or suspended sediments at levels that are higher in the lower river than in the upper river. The WQB results, which were obtained from large-volume samples, showed many more significant results. Because the data sets are so different, it is difficult to draw a general conclusion more specific than that some differences are detectable using the available data. Not only are the sampling strategies and analytical methods different, but also the substances analyzed. More differences might be detected if adequately sensitive methods of analysis were available. This was observed (1) in comparing the number of organic substances significantly different using the more sensitive WQB methods relative to those of MOE and (2) in the analysis of MOE transect data, where more differences were detectable for the two metals that had the highest concentrations.

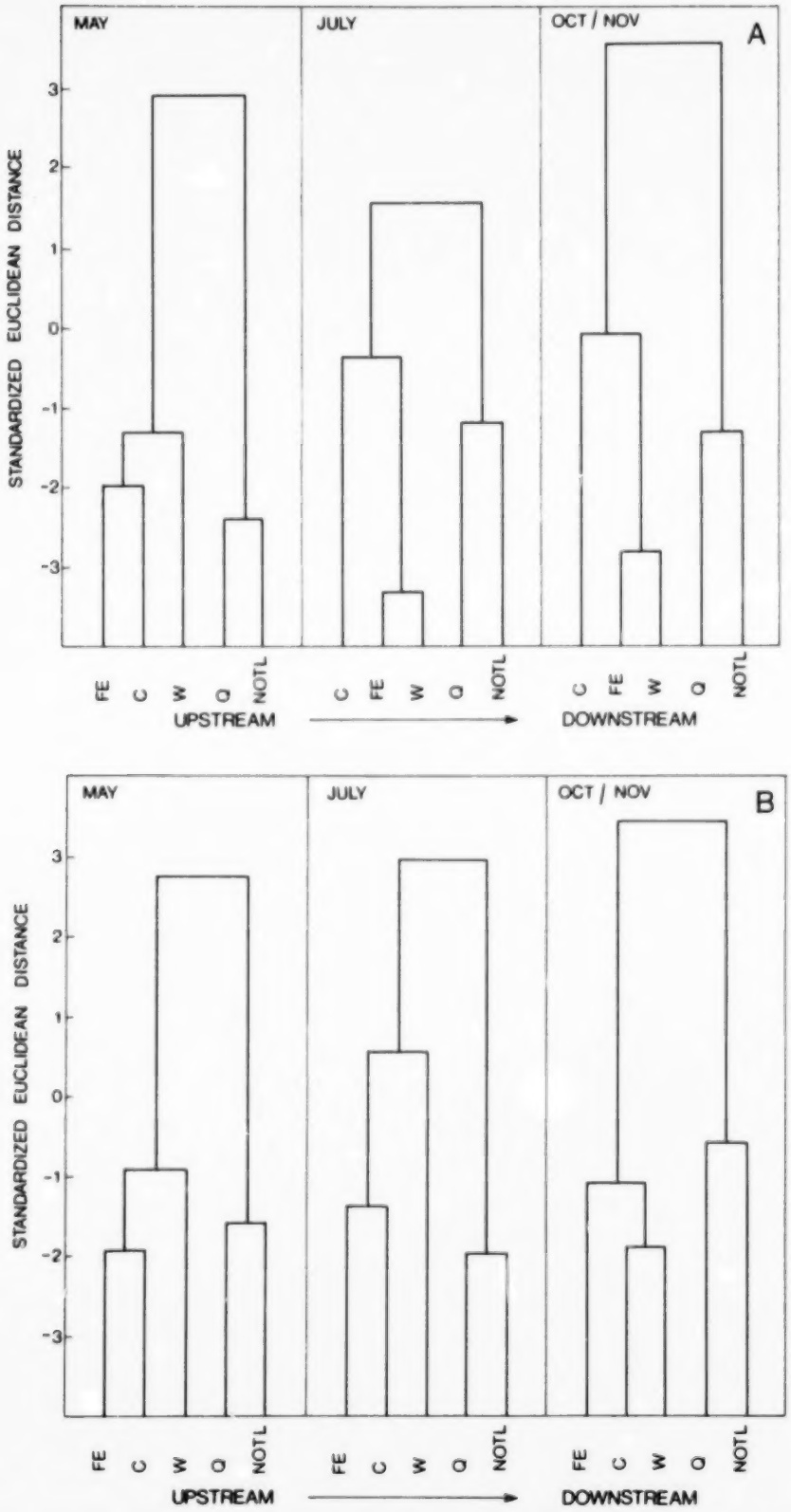

FIG. 4. Complete linkage clustering using the standardized Euclidean distance for (A) WQB suspended sediment data and (B) WQB aqueous phase data.

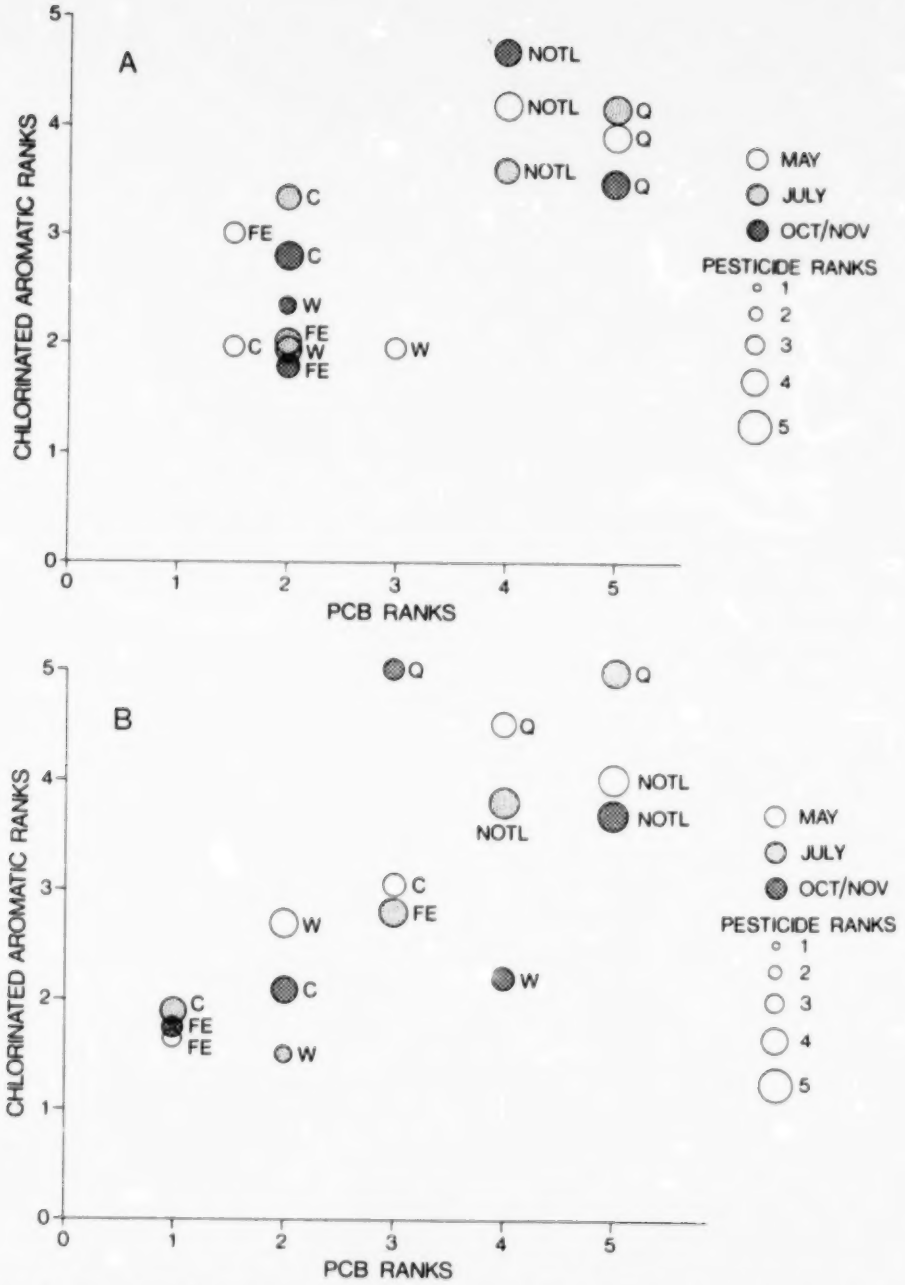

FIG. 5. Ranks for chlorinated aromatics, pesticides, and PCBs by sampling location and date for (A) WQB suspended sediment data and (B) WQB aqueous phase data.

Separate comments are appropriate for the individual data sets. The least powerful statistical method was used for the drinking water data, since this was the least quantitative of the three sets. Further, the samples were collected over a number of years, across seasons within a year, and over a long enough time to make changes in analytical methodology of concern. The sign test, as used here, is a simple procedure that can accommodate these limitations in the data. By comparing the level of a substance at NOTL and FE on matched sampling dates, the above-mentioned effects due to sampling over time can be eliminated at least to the extent possible without further information. Since two nonquantitative responses, trace (Tr) and not detected (ND), were recorded as well as concentrations, ranking of the outcome from a pair of stations as described above is a procedure that retains the most information without introducing further assumptions. The more powerful signed-rank test could be used if we could rank pairs such as the following: (Tr, ND) and (0.25, 0.32). This may be feasible for certain sets of data but it was not further considered here.

In summary, the two most important features of the analysis of the drinking water data are (1) the use of the information about when the samples were collected to eliminate time effects

TABLE 6. Statistics for the differences between locations for individual substances (*significant at 5%; **significant at 1%; ***significant at 0.1%; BD = binomially distributed)

Parameter	Suspended phase S_i	Aqueous phase S_i	Combined S_i
Total PCBs	157**	159**	316**
Pesticides			
Aldrin	BD	BD	BD
Dieldrin	131	136	267
A-BHC	154.5*	149	303.5**
G-BHC (Lindane)	BD	155**	—
A-Chlordane	133	136	269
G-Chlordane	BD	BD	BD
o,p-DDT	BD	BD	BD
p,p-DDT	BD	BD	BD
p,p-DDE	131	152*	284
p,p-DDD (TDE)	BD	BD	130.5
I-Endosulfan	145	141	286
II-Endosulfan	129	BD	—
Endrin	132.5	135	267.5
Heptachlor	BD	BD	BD
Heptachlor Epoxide	153*	95	248*
Methoxychlor	142	—	—
Mirex	157**	87	244*
Chlorinated aromatics			
1,2-Dichlorobenzene	152.5*	137	289.5
1,3)Dichlorobenzene	133	144	277
1,4-Dichlorobenzene	95	151*	246*
1,2,3-Trichlorobenzene	158**	159**	317***
1,2,4-Trichlorobenzene	155**	154*	309***
1,2,5-Trichlorobenzene	129.5	152*	281.5
1,2,3,4-Tetrachlorobenzene	157**	154*	311***
1,2,3,5- and 1,2,4,5-Tetrachlorobenzene	159**	155**	314***
Pentrachlorobenzene	157**	157**	314.5***
Hexachlorobenzene	156**	154*	310***
Metals			
Mercury	161**	—	—
Arsenic	143	—	—
Selenium	BD	—	—

and (2) the ranking of the outcome of a pair as a means of using the sign test when differences between quantitative levels cannot be calculated for all pairs. Application of the sign test, in a similar fashion, would often permit tests of significance to be performed where this is not presently being done. Since no assumptions about the distribution of the concentration of a substance need be made, if applied in keeping with the design of the data collection, the sign test should increase confidence in our conclusions.

Although the question of primary concern here is a difference in level between two regions of the Niagara River, a more extensive analysis might include the examination of the constancy of the difference over time. The plot of the cumulative sum for the number of pairs in which the NOTL level was above that of FE is the type of plot that can also be applied to ranks or quantitative data.

In contrast with the drinking water data, where useable records for individual substances were of varying length, sometimes providing many measurements for one substance, data were obtained at only three times in the WQB study, but results were reported for all substances each time. Thus, univariate procedures, used exclusively with the drinking water data, are supplemented by multivariate procedures that compare the results at different stations for all the substances simultaneously. Since the outcome was either ND or a concentration, more powerful rank procedures could be applied. For each substance, at any one sampling time, the five locations were ranked according to the level of the substance. The clustering of locations based on the Euclidean distance identified stations with the most similar ranks over all substances. The tests of the hypothesis of no difference in individual substances due to location were possible, since five locations were sampled even though only three sampling times were used. The use of these two types of methods together shows clearly which stations are different and for which substances.

Parametric procedures are preferable when distributional assumptions can be satisfied, and thus, normal-theory methods were used with the four tested metals. The size of the difference in concentration that can be measured is so small relative to the level of concentration for the other four metals that one would not expect to be able to detect differences between stations. If differences between means for each day are examined, month effects can be eliminated. Using the sign test with these differences, the significance levels for the one-sided test of the hypothesis of no difference are 0.94, 0.34, 0.81, and 0.88 for nickel, chromium, cadmium, and lead, respectively. Application of the more powerful Wilcoxon signed-rank test (Lehmann and D'Abrera 1975) is neither warranted, in view of the results of the sign test, nor practicable, due to ties, for these metals.

The total strength of the evidence of differences between the upper and lower Niagara River must be based on a synthesis of the results from individual data sets. A measure of uncertainty can be attached to the conclusions from the individual data sets but not to the conclusion based on all the data sets. This is not peculiar to the Niagara River data sets.

Acknowledgements

The statistical analysis of these data sets was requested by the Niagara River Toxics Committee. The authors would like to thank Mr. T. Wagner and Mr. D. Williams, Inland Waters Directorate – Ontario Region, for permission to publish the results and helpful discussion, and the staff at the Ontario Ministry of the Environment who provided the data, in particular Mr. P. B. Kauss. The helpful suggestions of the referees and editor were appreciated.

References

ALLAN, R. J., A. MUDROCH, AND M. MUNAWAR [ED.] 1983. The Niagara River – Lake Ontario pollution problem. J. Great Lakes Res. 9: 232 p.

ANDERBERG, M. R. 1973. Cluster analysis for applications. Academic Press, New York and London. 359 p.

BENNETT, C. A., AND N. L. FRANKLIN. 1954. Statistical analysis in chemistry and the chemical industry. John Wiley & Sons, Inc., New York, London, Sydney. 724 p.

BROWNLEE, K. A. 1965. Statistical theory and methodology in science and engineering. John Wiley & Sons, Inc., New York, London, Sydney. 590 p.

DATA QUALITY SUBCOMMITTEE. 1984. Final report of DQS to Niagara River Toxics Committee. Inland Waters Directorate, Ontario Region, Water Quality Branch, Burlington, Ont.

DRAPER, N. R., AND H. SMITH. 1981. Applied regression analysis. John Wiley & Sons, Inc., New York, Chichester, Brisbane, Toronto, Singapore. 709 p.

ENVIRONMENT CANADA. 1979. Analytical methods manual. Water Quality Branch, Inland Waters Directorate, Ottawa, Ont.

HOLLANDER, M., AND D. A. WOLFE. 1973. Nonparametric statistical methods. John Wiley & Sons, Inc., New York, London, Sydney, Toronto. 503 p.

KENDALL, M. G., AND A. STUART. 1968. The advanced theory of statistics. Vol. 3. Design and analysis, and time series. Hafner Publishing Company, New York, NY. 557 p.

KUNTZ, K. W., AND N. D. WARRY. 1983. Chlorinated organic contaminants in water and suspended sediments of the Lower Niagara River. J. Great Lakes Res. 9(2): 241–248.

LEHMANN, E. L., AND H. J. M. D'ABRERA. 1975. Nonparametrics, statistical methods based on ranks. Holden Day Inc., San Francisco, CA. 457 p.

MCCREA, R. C. 1982. Development of an aqueous phase liquid – liquid extractor (APLE). Interim Report. Inland Waters Directorate, Ontario Region, Water Quality Branch, Burlington, Ont.

ONTARIO MINISTRY OF THE ENVIRONMENT (MOE). 1984. Analytical methods outline handbook. Laboratory Service Board, Toronto, Ont.

PUTTER, J. 1955. The treatment of ties in some nonparametric tests. Ann. Math. Stat. 26: 368–386.

SCHEFFÉ, H. 1959. The analysis of variance. John Wiley & Sons, Inc., Chichester, Brisbane, Toronto, Singapore. 477 p.